PDMI PREPRINT — 26/1998

The Plancherel Measure of the Young-Fibonacci Graph

Alexander GNEDIN and Sergei KEROV

Institute of Mathematical Stochastics
University of Göttingen
Lotzestrasse 13, 37083 Göttingen, Germany
Russian Academy of Sciences
Steklov Mathematical Institute
St. Petersburg Department
Fontanka 27, St. Petersburg 191011, Russia
E-mail: gnedin@math.uni-goettingen.de, kerov@pdmi.ras.ru

November 1998

ABSTRACT

The Young-Fibonacci graph $\mathbb{YF}$ is an important example (along with the Young lattice) of differential posets studied by S. Fomin and R. Stanley. For every differential poset there is a distinguished central measure called the Plancherel measure. We study the Plancherel measure and the associated Markov chain, the Plancherel process, on the Young-Fibonacci graph.

We establish a law of large numbers which implies that the Plancherel measure cannot be represented as a nontrivial mixture of central measures, i.e. is ergodic. Our second result claims the convergence of the level distributions of the Plancherel measure to the GEM(1/2) probability law in the space of nonnegative series with unit sum, which is a particular example of distribution from the class of Residual Allocation Models.

In order to obtain the Plancherel process as an image of a sequence of independent uniformly distributed random variables, we establish a new version of the Robinson-Schensted type correspondence between permutations and pairs of paths in the Young-Fibonacci graph. This correspondence is used to demonstrate a recurrence property of the Plancherel process.

The second author partially supported by the Grant RFBR–NNIO 96-01-00096.

1. Introduction

In this paper we study the Plancherel measure of the Young-Fibonacci modular lattice. The Young-Fibonacci lattice provides a distinguished example of a graded graph, and the Plancherel measure is a distinguished central measure on the path space of this graph.

Graded graphs (also called Bratteli diagrams) and central measures arise in various contexts. For instance, for every increasing sequence of finite groups $G_1 \subset G_2 \subset \dots$ there is a graded graph describing the branching of characters of these groups. The central measures of this graph are in a bijective correspondence with the characters of the limiting group $G_\infty = \varinjlim G_n$.

An infinite oriented graph Γ is called *graded* if it has the following properties:

(1) there is a single source vertex $\varnothing$ (called the *root*) with no incoming edges;
(2) there is at least one outgoing edge for every vertex of the graph;
(3) all oriented paths from the root to a given vertex v have the same length called the *rank* of v.

Let Γ_n be the subset of vertices of rank n. We assume that the *levels* Γ_n are finite. If there is an edge leading from v to w, we write $v \nearrow w$.

Consider a Markov chain on Γ starting from the root, determined by transition probabilities $p(v,w)$ along the directed edges $v \nearrow w$ of the graph. Given a finite path $t = (v(0), v(1), \dots, v(n))$ of length n, we set $p(t) = \prod_{k=1}^{n} p(v(k-1), v(k))$. The Markov measure is *central* if $p(t)$ depends on the last vertex $v(n)$ of the path t only, for all n.

There is a general construction associating central measures to some infinite *regular* paths in Γ. The central measures which can be obtained by this method form the *Martin boundary* of the graph. Other central measures can be represented as mixtures of measures from the Martin boundary. A central measure is *ergodic* if it cannot be represented as a mixture of other central measures. All ergodic measures belong to the Martin boundary of the graph and form the *minimal boundary*.

The Young-Fibonacci graph $\mathbb{YF}$ (see a few levels on Fig. 1) was introduced by S. Fomin [3] and R. Stanley [9] as an important example of what they called *dual graphs* or *differential posets*. Another classical example is the Young graph (which has Young diagrams as vertices) arising in the representation theory of symmetric groups. One of the general features of differential posets is the existence of a remarkable central measure, called the *Plancherel measure*. In case of the Young graph the Plancherel measure corresponds to the regular representations of symmetric groups. See [7] for the combinatorial theory of locally semisimple algebras, and for more details on the boundary of the Young lattice.

The Young-Fibonacci graph describes the branching of representations for a family of finite dimensional semisimple algebras introduced by S. Okada [8] in terms of generators and relations. The central measures on the graph correspond bijectively to the characters of the limiting Okada algebra.

The Martin boundary of the graph $\mathbb{YF}$ was recently described by F. Goodman and the second author [5]. In particular, it was shown that the Plancherel measure belongs to the Martin boundary. The problem of identification of ergodic central measures of the graph $\mathbb{YF}$ was not solved in [5].

We start in Sections 2 and 3 with a detailed description of the Young-Fibonacci graph and its Plancherel measure. Then we show in Section 4 that the Plancherel measure of the Young-Fibonacci graph is ergodic. The proof is based on an appropriate law of large numbers for the Plancherel measure.

In Section 5 we represent the vertices of the Young-Fibonacci graph as the points of the simplex Δ of nonnegative series with unit sum. Then we show that the level distributions M_n

of the Plancherel measure converge in Δ to the distribution GEM(1/2), a particular case of the celebrated Griffiths-Engen-McCloskey distributions (see [6] for the history and applications of GEM distributions). In particular we show that a typical big word with respect to the Plancherel measure is made of large blocks of 2's separated by occasional occurrences of 1's.

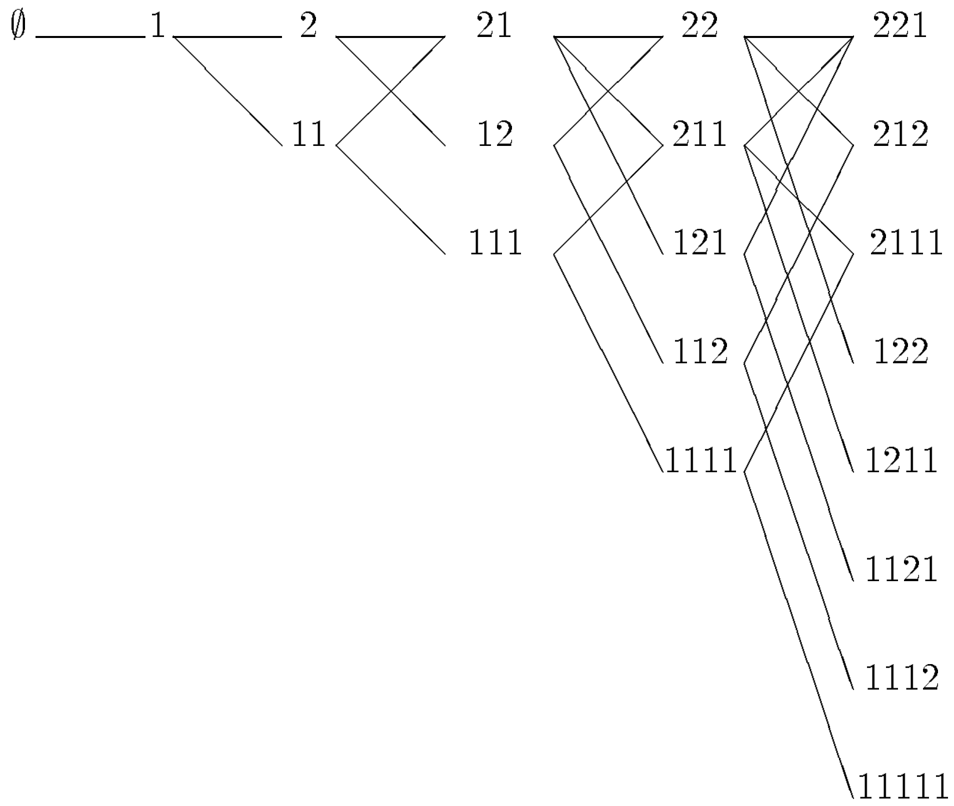

Fig. 1 The Young-Fibonacci graph.

In Section 6 we introduce an algorithm, called *Fibonacci solitaire*, which can be applied to a sequence of independent real random variables with a common continuous distribution, and results in the Plancherel process on $\mathbb{YF}$. The properties of our algorithm are similar to those of the generalized Robinson-Schensted correspondence defined for $\mathbb{YF}$ by S. Fomin [4]. Nevertheless, the two algorithms are different and no direct relations between them are known.

Finally, we use in Section 7 the Fibonacci solitaire to justify a remarkable recurrent behavior of the Plancherel process similar to that of the simple random walk on integers.

Acknowledgements. This piece of work was supported by the RFBR-NNIO 96-01-00096 grant and by the University of Göttingen. The second author greatly appreciates the hospitality of the Institute of Mathematical Stochastics, University of Göttingen, where the paper was completed. The authors thank Sergey Fomin for his inspiring comments on the combinatorics of Fibonacci solitaire. The proof of Theorem 6.2 is based on these comments.

2. The Young-Fibonacci graph

Let $\mathbb{YF}$ be the set of all finite sequences (words) of digits 1 and 2. The sum of digits $|v| = j_1 + \ldots + j_k$ of a word $v = j_1, \ldots, j_k$ will be referred to as the *rank* of v. For every $n = 0, 1, \ldots$ we denote by $\mathbb{YF}_n$ the subset of words of rank n. In particular, the empty word $\varnothing$ is the only element of $\mathbb{YF}_0$. Clearly, the set $\mathbb{YF} = \bigcup \mathbb{YF}_n$ is the disjoint union of $\mathbb{YF}_n$'s which we call *levels*.

We endow the set $\mathbb{YF}$ with the structure of a graded graph as follows. Given a word $v \in \mathbb{YF}$, we say that the maximal block of digits 2 at the left end of v is the *head* of v. We define the *head length* $h(v)$ as the number of digits in the head. For instance, the word $v = 22121$ has the head length $h(v) = 2$, and the head of $u = 1222$ is empty, so that $h(u) = 0$. For $0 \leq k \leq h(v)$ denote by v_k the word obtained from v by inserting 1 as the kth digit, reading left to right. If v contains at least one digit 1, we denote by v_* the word obtained from v by replacing the leftmost occurrence of 1 by 2. We call the words $v_0, v_1, \ldots, v_h, v_*$ the *successors* of v. For example, the word $v = 22121$ has the successors $v_0 = \mathbf{1}22121$, $v_1 = 2\mathbf{1}2121$, $v_2 = 22\mathbf{1}121$ and $v_* = 22\mathbf{2}21$.

Note that the rank of every successor of v is one bigger than the rank of the word v. If w is a successor of v, we write $v \nearrow w$. This relation determines the structure of an oriented graph on the set $\mathbb{YF}$ (see Fig. 1). Clearly, the length of any directed path from $\varnothing$ to v is equal to the rank of v.

The graph $\mathbb{YF}$ can be considered as the Hasse diagram of a partially ordered set with the minimal element $\varnothing$. It is often called the *Young-Fibonacci lattice* (this poset is indeed a modular lattice).

3. The Plancherel measure

We introduce in this section the central object of our study: a probability measure M on the path space of the graph $\mathbb{YF}$.

A finite path in the graph $\mathbb{YF}$ is a sequence of words

$$t = (v(m), v(m+1), \ldots, v(n)); \qquad v(k) \in \mathbb{YF}_k,$$

such that $v(j-1) \nearrow v(j)$ for every $j = m+1, \ldots, n$. We say that $v(m)$ is the initial word, and that $v(n)$ is the final word of the path t. We shall also consider infinite paths

$$t = (v(m), v(m+1), \ldots).$$

Let T denote the set of all infinite paths starting with the empty word $v(0) = \varnothing$. We call T the *path space* of the graph $\mathbb{YF}$. We endow T with the product topology, so that T is compact and totally disconnected.

Given a word v of rank $n = |v|$ and one of its successors w, we define a positive number $p(v, w)$ by setting

$$p(v, v_k) = \frac{n}{n+1} \cdot \ldots \cdot \frac{n-2k+2}{n-2k+3} \cdot \frac{1}{n-2k+1} \tag{3.1}$$

if $w = v_k$ for some k, $0 \leq k \leq h(v) = h$, and

$$p(v, v_*) = \frac{n}{n+1} \cdot \ldots \cdot \frac{n-2h+2}{n-2h+3} \cdot \frac{n-2h}{n-2h+1} \tag{3.2}$$

if $w = v_*$.

The following lemma shows that we can take the numbers $p(v, w)$ as transition probabilities of a Markov chain.

Lemma 3.1. *For every word v the numbers $p(v, w)$ sum up to unity,*

$$\sum_{w : v \nearrow w} p(v, w) = 1; \qquad v \in \mathbb{YF}. \tag{3.3}$$

Proof. Let $h = h(v)$ be the head length of v, and consider a sequence of independent events A_j, $j = 0, 1, \dots, h$ with probabilities $P(A_j) = 1/(n - 2j + 1)$. Clearly, (3.1) provides the probability that the first event to occur is A_k, while (3.2) is the probability that none of them occur. Therefore, the formula (3.3) holds true. $\square$

Definition. The Markov chain with the state space $\mathbb{YF}$, the empty word $\varnothing$ as the initial state, and transition probabilities $p(v, w)$, $v \nearrow w$, is called the *Plancherel process.* The probability law of the chain is called the *Plancherel measure* on T, and denoted M. The reason for this name is the analogy with the Plancherel measure on the Young lattice (cf. [7]).

In order to describe the probabilities of cylinders for the measure M we need further notations. Consider an entry of the digit 2 in a word v, and denote by u and w the subwords to the left and to the right of this entry, so that $v = u2w$. Following [5], we associate with this entry its *position* $d = |w| + 1$. For instance, the positions of 2's in the word 22121 are $7, 5, 2$ (reading left to right). Denote by $D(v)$ the number of directed paths from the empty word $\varnothing$ to a word v of rank n. It is well known that

$$D(v) = \prod_i d_i(v) \tag{3.4}$$

where the product is over all occurrences of the digit 2 in the word v. For instance, there are 70 paths with the final word $v = 22121$, so that $D(v) = 70$. For the sake of completeness we include a short proof of (3.4) communicated to the authors by Sergey Fomin.

The argument is by induction on the rank. Any word v can be written either as $1w$ or as $2w$, and we assume that $D(w) = \prod d_i(w)$ holds for every w of rank n.

If $v = 1w$ then v has only one immediate predecessor w, so that $D(v) = D(w)$ and (3.4) holds also for v.

The case $v = 2w$ is more delicate. Let $t = \big(w(0), w(1), w(2), \dots, w(n)\big)$ be a path in $\mathbb{YF}$ with the final word $w(n) = w$. Note that the position of the leftmost 2 in v is $d = n + 1$. Given an integer k in the range $1 \le k \le d$, consider the sequence of words

$$\big(w(0), w(1), \dots, w(k-1), 1w(k-1), 2w(k-1), 2w(k), \dots, 2w(n)\big) \tag{3.5}$$

obtained from t by inserting two words $1w(k-1)$, $2w(k-1)$ right after the word $w(k-1)$, and by replacing the words $w(j)$, $j \ge k$, by the corresponding words $2w(j)$. For instance, $t = (\varnothing, 1, 11, 21, 121, 221)$ and $k = 3$ produce $(\varnothing, 1, 11, 111, 211, 221, 2121, 2221)$. Clearly, the sequence (3.5) is a path in $\mathbb{YF}$ with the final word $v = 2w$, and every such path arises from a specific pair (k, t), Therefore, $D(2w) = d\,D(w)$ and (3.4) follows from the inductive hypothesis.

Given a finite path $s = (v(0), v(1), \dots, v(n))$, we denote by T_s the cylinder set of infinite paths in T, such that their initial part coincides with s. There is a simple formula for the cylinder probabilities in terms of the function D.

Lemma 3.2. *For every path $s = (v(0), \dots, v(n))$ with the final vertex $v(n) = v$,*

$$M(T_s) = \frac{D(v)}{n!}. \tag{3.6}$$

Proof. Consider a path s of length n, and a longer path $t = (v(0), \dots, v, w)$ with the final word $w = v_k$. We have

$$\frac{D(v_k)}{(n+1)!}\frac{n!}{D(v)} = \frac{1}{n+1}\prod_{i=1}^{k}\frac{d_i+1}{d_i} = \frac{1}{n+1}\prod_{i=1}^{k}\frac{n-2i+2}{n-2i+1} = p(v,w) = \frac{M(T_t)}{M(T_s)}.$$

A similar conclusion holds if the final word of the path t is $w = v_*$. Assuming that the formula is true for every path of length n we conclude that it is true for the paths of length $n+1$. □

Note that the cylinder probabilities $M(T_s) = D(v)/n!$ depend solely on the final word v of the path s. The probability measures on T with this property are called *central* [7].

We denote by $V_n(t)$ the nth word $v(n)$ in a path $t = (v(0), v(1), \dots, v(n), \dots)$ and we regard V_n as a random variable on the probability space (T, M). Thus, the random sequence (V_n) is a realization of the Plancherel process. Denote by $M_n(v) = M\{V_n = v\}$ the probability to cross the nth level $\mathbb{YF}_n$ in a particular word $v \in \mathbb{YF}_n$, i.e., M_n is the distribution of V_n. It follows from (3.4) and (3.6) that

$$M_n(v) = \frac{D^2(v)}{n!}. \tag{3.7}$$

Summing up these probabilities we arrive at a fundamental identity of the theory of differential posets,

$$\sum_{v \in \mathbb{YF}_n} D^2(v) = n!. \tag{3.8}$$

In case of the Young lattice, this formula translates as the Burnside identity for the symmetric group.

4. Ergodicity of the Plancherel measure

A central probability measure N on the path space T is *ergodic* if it cannot be represented as a nontrivial mixture of central measures. Equivalently, the zero-one law must hold: $N(A) = 0$ or 1 for any *tail event* $A \subset T$. Recall that A is a tail event if for any $t \in A$ we have $s \in A$ for all paths $s \in T$ coinciding eventually with t. Each central measure can be represented as a mixture of ergodic measures.

A common way to obtain central measures is via the following construction. Let $D(w, v)$ be the number of directed paths with initial word w and final word v. In particular, $D(\varnothing, v) = D(v)$ is the number of paths from the root to v. Let $s = (w(0), w(1), \dots, w(m))$ be a path with the final word $w(m) = w$. An infinite path $t = (v(0), v(1), \dots, v(n), \dots)$ is said to be *regular*, if the limit in the right hand side of

$$N(T_s) = \lim_{n \to \infty} \frac{D(w, v(n))}{D(v(n))}$$

exists for all $w \in \mathbb{YF}$. We take this limit as a definition of the cylinder probability $N(T_s)$ of a central measure N. We say that t *generates* the measure N. The set of measures obtained by this construction is called the *Martin boundary.* All ergodic central measures belong to the Martin boundary.

The following general facts follow from [1], Theorem 1.1, or can be derived from a version of Birkhoff Ergodic Theorem, cf. [7]:

(a) for any central measure N the set of regular paths is of full N-measure;

(b) for N ergodic, the set of regular paths generating N is of full N-measure.

Specifically for the Young-Fibonacci graph, the explicit regularity conditions were found in [5] and involve the function $\pi : \mathbb{YF} \to [0, 1]$ defined as

$$\pi(v) = \prod_{i : d_i > 1} \left(1 - \frac{1}{d_i}\right). \tag{4.1}$$

The product in (4.1) is taken over all positions $d_i > 1$ of 2's in v. More precisely, it was shown in [5], Theorem 8.5, that

(c) $\pi(v(n))$ converge to a limit along each regular path $t = (v(0), v(1), \ldots, v(n), \ldots)$;

(d) $\pi(v(n)) \to 0$ if and only if t generates the Plancherel measure.

Taken together, (a) and (c) imply that $\pi(v(n))$ converge N-almost surely for any central measure N. In order to prove the ergodicity of the Plancherel measure M we shall show that $\pi(v(n)) \to 0$ a.s. with respect to M.

Proposition 4.1. *As $n \to \infty$, the random variables $\pi_n = \pi(V_n)$ converge to zero M-almost surely.*

Proof. Consider

$$m_n = \sum_{v \in \mathbb{YF}_n} \pi(v) M_n(v), \tag{4.2}$$

the mean value of the variable π_n on the space T. We wish to show that for $n = 2k+1$ or $n = 2k+2$

$$m_n = \frac{(2k)!!}{(2k+1)!!}. \tag{4.3}$$

To this end, we derive the recurrence relation

$$m_{n+1} = \frac{1}{n+1} m_n + \frac{n-1}{n+1} m_{n-1}, \tag{4.4}$$

valid for every $n = 1, 2, \ldots$. Indeed, a word $w \in \mathbb{YF}_{n+1}$ can be written either as $w = 1v$ with $v \in \mathbb{YF}_n$, or as $w = 2u$ with $u \in \mathbb{YF}_{n-1}$. It follows readily from (3.4) and (3.7) that in the first case

$$\pi(1v) = \pi(v); \qquad M_{n+1}(1v) = \frac{M_n(v)}{n+1},$$

while in the second case

$$\pi(2u) = \left(1 - \frac{1}{n}\right)\pi(u); \qquad M_{n+1}(2u) = \frac{n\, M_{n-1}(u)}{n+1}.$$

Now (4.4) follows directly from these calculations and definition (4.2). Equation (4.3) follows by observing that $m_1 = m_2 = 1$ and the right-hand side satisfies the recurrence relation (4.4).

It is not hard to obtain from (4.3) that $m_n \sim const/\sqrt{n}$ and thus $m_n \to 0$. Since the random variables π_n are nonnegative, we can conclude that π_n converge to 0 in mean, hence also in probability. By a standard argument, π_n go to zero almost surely at least along a subsequence (n_k). But then the existence of the limit implies that the whole sequence (π_n) goes to zero. $\square$

We can now prove one of our main results.

Theorem 4.2. *The Plancherel measure M on the path space T is ergodic.*

Proof. We just need to combine what we already know. For every ergodic central measure N, the sequence $\pi(v(n))$ converges to a constant, say c_N, along N-almost all paths $(v(0), v(1), \ldots)$. In fact, the limit of $\pi(v(n))$ is measurable with respect to the sigma algebra of tail events, and the latter is trivial in the ergodic case by definition. Furthermore, for $N \neq M$ we have $c_N > 0$ in view of (b) and (d). On the other hand, Proposition 4.1 says that the limit for M is zero. Thus the Plancherel measure cannot be represented as a mixture of ergodic measures $N \neq M$, and is therefore ergodic. $\square$

It was conjectured in [5] that *all* measures in the Martin boundary of the Young-Fibonacci graph are in fact ergodic. Apart from the Plancherel measure and a minor class of finitely supported central measures, the ergodicity conjecture remains open.

5. Weak convergence of random words V_n

In this Section we establish a limit theorem for the distributions M_n of random words V_n. To this end, we embed the level sets $\mathbb{YF}_n$ of the Young-Fibonacci graph into the simplex Δ of all sequences

$$\mathbf{x} = (x_1, x_2, \dots, x_n, \dots)$$

such that $\sum_j x_j \leq 1$. Then we show that the images of the M_n's in Δ weakly converge to the GEM(1/2) distribution.

Consider a word $v \in \mathbb{YF}_n$ as a sequence of contiguous blocks of digits 2 separated by 1's. Each block contains at most one digit 1 which is its terminating digit. The rightmost block may contain no 1 at all. Denote by $r_1, r_2, \dots$ the sequence of ranks of blocks, reading left to right, so that $r_1 + r_2 + \ldots = n$. In order to obtain a point in Δ the sequence should be complemented by zeros. For instance, if $v = 122112$ then $r_1 = 1$, $r_2 = 5$, $r_3 = 1$, $r_4 = 2$ and $r_j = 0$ for $j \geq 5$. We define a point $\mathbf{x}(v) \in \Delta$ as

$$\mathbf{x}(v) = \Big(\frac{r_1(v)}{n}, \frac{r_2(v)}{n}, \dots\Big) \tag{5.1}$$

where $n = |v|$ is the rank of the word v. We identify the distribution M_n of the random word V_n with its image in Δ under the map $\mathbf{x} : \mathbb{YF}_n \to \Delta$ defined by (5.1).

GEM(θ) probability law on Δ appears in connection with a Residual Allocation Model

$$\mathbf{X} = (X_1, X_2, \dots, X_n, \dots) \in \Delta \tag{5.2}$$

where

$$X_1 = U_1, \qquad X_n = (1 - U_1)(1 - U_2)\dots(1 - U_{n-1})U_n, \qquad n = 2, 3, \dots \tag{5.2}$$

and $U_1, U_2, \dots$ are independent identically distributed random variables with the beta density $\varphi_\theta(u) = \theta(1-u)^{\theta-1}$ on the interval $[0, 1]$ (see [6]).

The variables X_n are called *residuals* and U_n *residual fractions*, since

$$U_n = X_n \big/ \big(1 - X_1 - \ldots - X_{n-1}\big). \tag{5.4}$$

Theorem 5.1. *The image in Δ of the distribution M_n of the random word V_n converges weakly to GEM(1/2) distribution, as $n \to \infty$.*

Proof. Focus on the first coordinate $r_1(V_n)/n$. We intend to show that it converges in distribution to the first coordinate X_1 of the sequence (5.2). Denote by $P_n(m)$ the probability that a random word V_n starts with at least m digits 2. It follows from (3.8) that

$$P_n(m) = \sum_{u:|u|=n-2m} \frac{D^2(2\dots 2u)}{n!} =$$
$$= \frac{(n-2m)!}{n!}\prod_{k=1}^{m}(n-2k+1)^2 \sum_{u:|u|=n-2m}\frac{D^2(u)}{(n-2m)!} =$$
$$= \frac{(n-2m)!}{n!}\prod_{k=1}^{m}(n-2k+1)^2 = \prod_{k=1}^{m}\frac{n-2k+1}{n-2k+2}.$$

Therefore, as $n\to\infty$

$$M_n\left\{\frac{r_1(V_n)}{n}\geq y\right\} = P_n(m) = \frac{\Gamma(n/2+1/2)}{\Gamma(n/2+1)}\frac{\Gamma(n/2-m+1)}{\Gamma(n/2-m+1/2)} \to (1-y)^{1/2}$$

where m is selected so that $2m+1$ is the largest odd integer $\leq ny$, and we have used the well known approximation

$$\frac{\Gamma(n+a)}{\Gamma(n+b)} \sim n^{a-b}; \qquad n\to\infty.$$

Since $1-(1-y)^{1/2}$ is the distribution function for the required beta law, we conclude that $r_1(V_n)/n \to X_1$ in distribution.

Now consider the distribution M_{m+n} on the level set $\mathbb{YF}_{m+n}$ and the subset of words $v=uw$ with a fixed initial part u of rank m. We show that the conditional distribution of the tail w, given u, equals M_n and does not depend on the choice of $u\in\mathbb{YF}_m$. Indeed, let $D_n(u)=\prod_i d_i$ be the product of positions of 2's in $v=uw$ belonging to the part u. Then

$$M_{m+n}(uw) = \frac{D^2(uw)}{(m+n)!} = \frac{n!D_n^2(u)}{(m+n)!}\frac{D^2(w)}{n!} = \frac{n!D_n^2(u)}{(m+n)!}M_n(w),$$

hence

$$\frac{M_{m+n}(uw)}{M_{m+n}\{uz:\ z\in\mathbb{YF}_n\}} = M_n(w) \tag{5.5}$$

and we are done.

We proceed with the proof that the variables $r_1(V_n)/n$ and $r_2(V_n)/(n-r_1(V_n))$ are asymptotically independent. To this end, we show that for every $t_1,t_2\in[0,1]$

$$M_n\left\{\frac{r_1(V_n)}{n}\leq t_1,\ \frac{r_2(V_n)}{(n-r_1(V_n))}\leq t_2\right\} \to \left(1-(1-t_1)^{1/2}\right)\left(1-(1-t_2)^{1/2}\right) \tag{5.6}$$

as $n\to\infty$. By (5.5), the probability in the left-hand side can be written in the form

$$\sum_{k\leq nt_1} M_n\{r_2(V_n)\leq (n-k)t_2 \mid r_1(V_n)=k\}\, M_n\{r_1(V_n)=k\} =$$
$$= \sum_{k\leq nt_1} M_{n-k}\{r_1(V_{n-k})\leq (n-k)t_2\}\, M_n\{r_1(V_n)=k\}.$$

We derive from the first part of the proof that

$$M_{n-k}\{r_1(V_{n-k}) \le (n-k)t_2\} \to \big(1-(1-t_2)^{1/2}\big)$$

as $n \to \infty$, and the limit does not depend on k. Since $n-k \ge (1-t_1)n \to \infty$, the convergence is uniform in k for all terms in the sum. In a similar manner we obtain that

$$\sum_{k \le nt_1} M_n\{r_1(V_n)=k\} = M_n\{r_1(V_n) \le nt_1\} \to \big(1-(1-t_1)^{1/2}\big)$$

as $n \to \infty$ and (5.6) follows.

Proceeding as above, we see that the vector

$$\left(\frac{r_1(V_n)}{n}, \frac{r_2(V_n)}{n-r_1(V_n)}, \dots, \frac{r_m(V_n)}{n-r_1(V_n)-\dots-r_{m-1}(V_n)}\right)$$

converges in distribution to the vector $(U_1, U_2, \dots, U_m)$ for every m. $\square$

Note that for almost all paths $t \in T$ the variables $r_1(V_n(t))/n$ have no limit. Therefore the vectors $\mathbf{x}(V_n)$ do not converge in the strong sense.

6. Fibonacci solitaire

We introduce in this Section a sort of solitaire game. More precisely, we describe an algorithm which can be applied to a permutation and produces a finite path in the Young-Fibonacci graph. The important property of the algorithm is that it transforms a uniformly distributed permutation into the Plancherel process.

Denote by S_n the set of integer sequences $s=(m_1, m_2, \dots, m_n)$, such that $1 \le m_k \le k$ for all $k=1,2,\dots,n$. Given a permutation $\sigma=(i_1, i_2, \dots, i_n)$ of $\{1,\dots,n\}$, we set $m_k = \#\{j \le k : i_j \le i_k\}$ for $k=1,2,\dots,n$. The vector $(m_1, m_2, \dots, m_n) \in S_n$ is called the *code* of the permutation. If σ is a random permutation with uniform distribution, then m_k is uniform on the set $\{1,2,\dots,k\}$, and $m_1, m_2, \dots, m_n$ are independent.

The core of our algorithm is an insertion procedure which applies to a word a in the alphabet $\{0,1,2\}$, and a given position k to the immediate right of the first k letters in a. We denote by n the number of letters in a, so that $0 \le k \le n$. The output is a word b with $n+1$ letters in the alphabet $\{0,1,2\}$. The procedure includes two cases.

Fibonacci Insertion Procedure.

Input: a word a of n letters in the alphabet $\{0,1,2\}$ and an integer $0 \le k \le n$.

Output: a word b of $n+1$ letters in the alphabet $\{0,1,2\}$.

(1) *If there are no digits 1 among the k leftmost letters of the word a, we obtain b by inserting 1 to the immediate right of the kth digit of a;*

(2) *if at least one of the first k digits of the word a is 1, we define b by inserting 0 right after the first k letters of a, and by replacing the leftmost occurrence of 1 in a by the letter 2.*

The output word is denoted as $b = FI(a,k)$.

For example, take $a = 22012001$. The position of the leftmost 1 in a is 4, hence for $k=0,1,2,3$ we apply Case 1 to obtain the words **1**22012001, 2**1**2012001, 22**1**012001 and 220**1**12001. For $k=4$ the output is $b=$ 220**20**2001. For $k=5,6,7$ we obtain equal words 220**220**001 $=$ 220**22**0**0**01 $=$ 220**2**200**0**1, and for $k=8$ the result is $b=$ 220**2**2001**0**.

Lemma 6.1. *Let $b = FI(a, k)$ for some k. Denote by $v = F(a)$ and $w = F(b)$ the words obtained from a and b by removing all 0's. Then w is an immediate successor of v in the Young-Fibonacci graph $\mathbb{YF}$.*

Proof. Clear from the descriptions of the successors of v in $\mathbb{YF}$, see Section 2. □

Given $s = (m_1, \dots, m_n) \in S_n$, we define a sequence of words $a_0, a_1, a_2, \dots, a_n$ recurrently, starting with the empty word a_0 and setting $a_k = FI\big(a_{k-1}, m_k\big)$ for $k = 1, \dots, n$. We are interested in the sequence

$$F_n(s) = \big(F(a_0), F(a_1), \dots F(a_n)\big).$$

It follows from Lemma 6.1 that $F_n(s)$ is a path in the Young-Fibonacci graph. In a similar way one can define the map $F_\infty : S_\infty \to T$ from the space S_∞ of infinite sequences $s = (m_1, m_2, \dots)$, such that $1 \le m_k \le k$, to the path space T of $\mathbb{YF}$.

Theorem 6.2. *Assume that the entries $m_1, m_2, \dots$ of the sequence $s \in S_\infty$ are independent and that m_n is uniformly distributed on the set $\{1, 2, \dots, n\}$ for all natural n. Then the path $F_\infty(s)$ has the Plancherel distribution M in the path space T of the Young-Fibonacci graph.*

Proof. Consider a finite path $t = \big(v(0), v(1), \dots, v(n)\big)$ in the graph $\mathbb{YF}$ with the final word $v(n) = v$. We shall check that there are exactly $D(v)$ sequences $s \in S_n$, such that $F_n(s) = t$. Comparison with the formula (3.6) for the Plancherel measure would imply the Theorem.

Consider a word a in the alphabet $\{0, 1, 2\}$ which contains equal numbers of the letters 0 and 2. By a *rigging* of a we mean a bijection between the set of letters 2 and the set of letters 0 in a, such that the 0 corresponding to an occurrence of 2 is always to the right of this 2.

Note that in the second case of the Fibonacci insertion we replace some occurrence of 1 by 2, and insert 0 in some position to the right of the new digit 2. As a result, the final word $a = a_n$ in a sequence $a_0, a_1, \dots, a_n$ corresponding to $(m_1, \dots, m_n) \in S_n$ comes up with a canonical rigging. Vice versa, consider a path $t = \big(v(1), \dots, v(n-1), v\big)$ in $\mathbb{YF}$, and a rigged word a such that $F(a) = v$. One can easily see that there exists a unique sequence $s = (m_1, \dots, m_n)$ such that $F_n(s) = t$ and the rigging of a is the one induced by the Fibonacci insertion. Indeed, taking into account the last two words $v(n-1)$ and v in the path t we know if a_n is obtained from a_{n-1} by the first or by the second case of the Fibonacci Insertion Procedure. In the first case we also know the digit 1 to be erased from a_n to obtain a_{n-1}, and we recover m_n as the number of letters to the left of this 1. In the second case we know which 2 in a_n is replaced by 1 in a_{n-1}, and we use the rigging of a_n to find the 0 associated with this occurrence of 2 in a_n. Then m_n is recovered as the number of letters in a_n to the left of this 0. Other elements $m_{n-1}, \dots, m_1$ are recovered in a similar way.

We have proved in the preceding paragraph that the number of sequences $s \in S_n$, such that $F_n(s) = t$, equals the number of rigged words a, such that $F(a) = v$ where v is the final word of the path t. Now we check that the number of such rigged words is $\prod d_i$, where the product runs over all 2's in the word v. Denote by $d_1, \dots, d_k$ the positions of 2's in v, reading right to left. Then there are exactly d_1 choices for a 0 to the right of the rightmost 2 in v. After we have inserted a 0 into v at a chosen place, we consider the next occurrence of 2 and note that there are d_2 choices for the position of the associated 0 to the left of this 2. Using the induction in i we see that the number of riggings equals $D(v)$, and the Theorem follows. □

Fibonacci solitaire yields a mapping from permutations of n integers to $\mathbb{YF}_n$. Taking a random permutation and using the representation (5.1) for Fibonacci words we get a distribution on Δ. Theorem 5.1 says that this distribution is close to GEM(1/2). The appearing of the parameter value 1/2 in connection with permutations looks somewhat surprising, especially if we recall that partitioning random permutations into cycles is governed by GEM(1).

A familiar method to generate permutations is to rank real numbers. Consider a sequence (x_n) of distinct reals. By definition, the *rank* m_n of the entry x_n is the number of entries $x_1, \dots, x_{n-1}, x_n$ smaller or equal than x_n. It is well known that for random independent identically distributed x_n's without ties the ranks $m_1, m_2, \dots$ are independent, with m_n uniformly distributed on $\{1, 2, \dots, n\}$. We define an infinite path $t = \mathcal{F}((x_n))$ in the Young-Fibonacci graph as $\mathcal{F}((x_n)) = F_\infty(m_1, m_2, \dots)$, see an example in Figure 2.

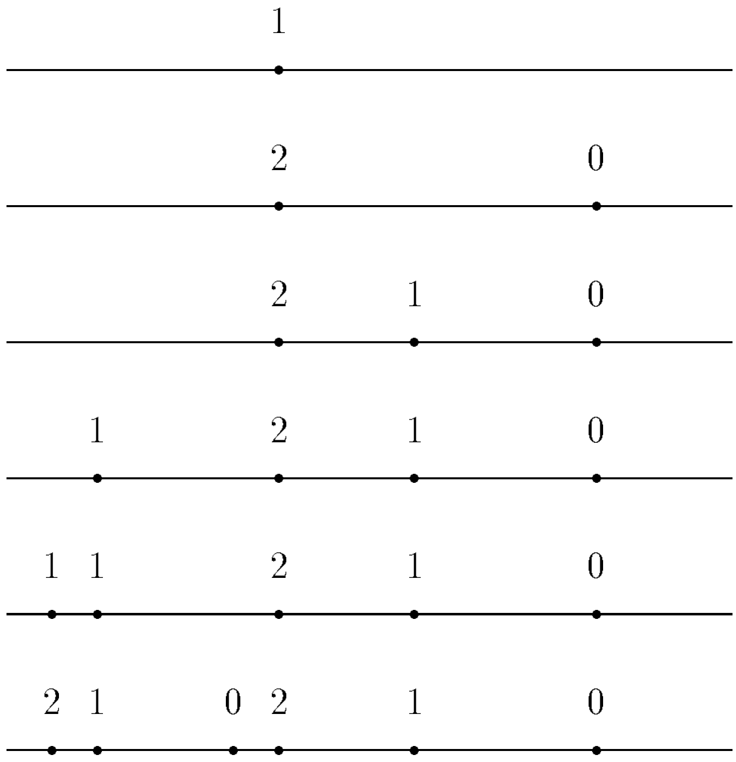

Fig. 2 Fibonacci solitaire.

Corollary 6.3. *Let (X_n) be a sequence of independent random real-valued random variables with a common continuous distribution function. Then the path $\mathcal{F}((X_n))$ is a realization of the Plancherel process on the graph $\mathbb{YF}$.*

More generally, one can take arbitrary exchangeable real random variables without ties to generate the Plancherel process via F_n.

The algorithm F_n of Fibonacci solitaire can be reformulated as a bijection between the set of permutations and the set of pairs of paths of length n with a common final word. Note that a similar bijection known as the Robinson-Schensted type correspondence for the Young-Fibonacci graph (cf. [4]) is different from ours.

7. Recurrence

The Plancherel process (V_n) is a transient Markov chain, because ranks increase steadily along the process. However, the chain demonstrates an interesting recurrent behavior related to visits in the set $\mathcal{T} = \{\varnothing, 2, 22, 222, \dots\}$.

Define $two(n) = 22\dots 2$ to be the word from $\mathbb{YF}_{2n}$ made of n digits 2. Let $E_{2n} = \{V_{2n} = two(n)\}$, where $n \geq 0$ and E_0 is a full event. The quantities of interest are

$$q_{2n} = M(E_{2n}); \qquad n \geq 0,$$

the probability to visit $\mathcal{T}$ on the $2n$th step, and

$$r_{2n} = M(E_0^c \cap E_2^c \cap \dots \cap E_{2(n-1)}^c \cap E_{2n}); \qquad n > 0$$

the probability to visit $\mathcal{T}$ at the $2n$th step for the first time.

Equations (3.4) and (3.7) imply

$$q_{2n} = \binom{2n}{n} 2^{-2n}, \tag{7.1}$$

where we recognize the return probability to zero for the simple random walk on integers (see [2]). Remarkably enough, this analogy goes further.

Introduce the first recurrence time

$$\tau_1 = \min\{2n > 0 : \quad V_{2n} \in \mathcal{T}\}$$

with the convention $\min \varnothing = \infty$ and inductively for $k = 2, 3, \dots$

$$\tau_k = \min\{2n > \tau_{k-1} : \quad V_{2n} \in \mathcal{T}\}.$$

Theorem 7.1. *All τ_k are finite $M-$almost surely, i.e. the Plancherel process visits the set $\mathcal{T}$ infinitely often. The lengths of excursions from $\mathcal{T}$*

$$\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_3, \dots$$

are independent identically distributed with

$$M(\tau_1 = 2n) = r_n = \frac{1}{n}\binom{2n-2}{n-1} 2^{-2n+1}. \tag{7.2}$$

Thus, the expected value of τ_1 as well as the expected lengths of excursions are infinite.

Proof. Fix n, and consider an infinite directed path t starting in $two(n)$. Each word on this path has n *old* digits 2 coming from $two(n)$ and some other digits 1 or 2. Removing all old 2's from each word along t results in a path from T, say t', which starts at the root. Furthermore, positions of the elements of $\mathcal{T}$ are the same for both paths t and t'. For example, for $t = (22, 221, 2121, 2221, 2222, \dots)$ we have $t' = (\varnothing, 1, 11, 21, 22, \dots)$, and these positions are $0, 4, \dots$.

The clue to the proof is the observation that taking for t the Plancherel process (V_n) and conditioning on the event $E_{2n} = \{V_{2n} = two(n)\}$ result in t' with the Plancherel distribution on $\mathbb{YF}$. This is best seen via the Fibonacci solitaire realization with i.i.d. variables, as in Corollary 6.3: E_{2n} is determined by $X_1, \dots, X_{2n}$ while t' depends only on ranking of $X_{2n+1}, X_{2n+2}, \dots$.

To induce t as in the above example we must have $X_6 < X_5$, $X_7 > X_6$, $X_8 > X_5$, ... This event is clearly independent of X_1, X_2, X_3, X_4.

We argue that visiting $\mathcal{T}$ is a periodic recurrent event, as in [2] Chapter 13, and the whole situation is well recognized. Indeed,

$$M(E_{2m}|E_{2n}) = M(E_{2(m-n)}), \qquad m \geq n,$$

and the events E_{2m} are conditionally independent of $E_0, \ldots, E_{2(n-1)}$ given E_{2n}. Conditioning on the first return of the Plancherel process in $\mathcal{T}$ we obtain

$$q_{2n} = r_2 q_{2n-2} + r_4 q_{2n-4} + \ldots + r_{2n} q_0,$$

which is a standard renewal equation. Recalling (7.1) and comparing with the classical setting we see that τ_1 has the same distribution as the recurrence time for the simple random walk, whence (7.2). In particular, $r_{2n} \sim 1/(2\sqrt{\pi} n^{3/2})$ and τ_1 has infinite expectation. $\square$

In other words, we obtain a null recurrent Markov chain if we modify the Plancherel process by requiring that the process returns from any state in $\mathcal{T}$ directly to $\varnothing$ (and keeping other transition probabilities). The analogy with random walk does not seem to go beyond the recurrence of $\mathcal{T}$: it breaks down completely if we consider returns to the set of words with a fixed positive number of 1's.

References

1. P. Diaconis and D. Freedman, *Partial exchangeability and sufficiency*, Statistics: Applications and New Directions (J.K. Ghosh and K. Roy, eds.), Indian Statistical Institute Calcutta, 1981, pp. 205–236.
2. W. Feller, *An Introduction to Probability Theory and its Applications*, vol. I, Wiley, NY, 1962.
3. S. V. Fomin, *Duality of graded graphs*, *J. Soviet Math.* 41 (1988), 979–991.
4. S. V. Fomin, *Schensted algorithms for dual graded graphs*, J. Algebraic Combinatorics **4** (1995), 5–45.
5. F. Goodman and S. Kerov, *The Martin Boundary of the Young-Fibonacci Lattice*, Accepted by the J. of Algebraic Combinatorics (available electronically via Math. Arch., Paper Nr. CO/9712266) (1997), 1–29.
6. N. L. Johnson, S. Kotz and N. Balakrishnan, *Multivariate Discrete Distributions*, Wiley, NY, 1997.
7. S. Kerov and A. Vershik, *The Grothendieck Group of the Infinite Symmetric Group and Symmetric Functions with the Elements of the K_0-functor theory of AF-algebras*, Representation of Lie groups and related topics (A. M. Vershik and D. P. Zhelobenko, eds.), Adv. Stud. Contemp. Math., vol. 7, Gordon and Breach, 1990, pp. 36–114.
8. S. Okada, *Algebras associated to the Young-Fibonacci lattice*, *Trans. Amer. Math. Soc.* 346 (1994), 549–568.
9. R. P. Stanley, *Differential Posets*, *J. Amer. Math. Soc.* 1 (1988), 919–961.